What is claimed is:

5    1.    A method for identifying, classifying or quantifying one or more nucleic acids in a sample comprising a plurality of nucleic acids having different nucleotide sequences, the method comprising:

(a) probing said sample with one or more recognition means wherein each recognition means recognizes a different target nucleotide subsequence or a different set of target nucleotide subsequences to provide one or more targeted nucleic acids;

10

(b) generating one or more first signals from said sample probed by said recognition means, each generated first signal arising from a targeted nucleic acid in said sample and comprising a representation of (*i*) the length between occurrences of target subsequences in said targeted nucleic acid, and (*ii*) the identities of said target

15    subsequences in said targeted nucleic acid or identities of said target subsequences among which are included the target subsequences in said targeted nucleic acid;

(c) selecting one or more targeted nucleic acids based on their corresponding first signals;

(d) extending sequence information from one or more target subsequences in said

20    selected targeted nucleic acid by one or more nucleotides providing one or more extended subsequences under conditions that generate one or more second signals arising from said selected targeted nucleic acid, at least one of whose subsequences has been extended, in said sample, wherein said second signal comprises a representation of (*i*) the length between occurrences of target subsequences, at least one of which has been extended, in

25    said nucleic acid, and (*ii*) the identities of said selected target subsequences, at least one of which has been extended, in said selected targeted nucleic acid or identities of said target subsequences, at least one of which has been extended, among which are included the target subsequences in said selected targeted nucleic acid; and

(e). searching a nucleotide sequence database to determine sequences that match or the

30    absence of any sequences that match one or more or of said selected targeted nucleic acids having at least one extended subsequence and represented by said generated second

31

signals, said database comprising a plurality of known nucleotide sequences of nucleic acids that may be present in the sample, wherein a sequence from said database is determined to match said selected targeted nucleic acid providing a generated second signal when the sequence from said database has both (*i*) the same length between

5      occurrences of target subsequences, at least one of which has been extended, as is represented by the generated signal, and (*ii*) the same target subsequences, at least one of which has been extended, as are represented by the generated signal, or target subsequences, at least one of which has been extended, that are members of the same sets of target subsequences represented by the generated signal,

10     whereby a matched nucleic acid in said sample is identified, classified, or quantified.

2.     The method of claim 1 wherein said second generated signal is a negative oligo-competition signal.

15     3.     The method of claim 1 wherein said second generated signal is a positive oligo-competition signal.

4.     The method of claim 2 wherein the extending of the sequence information comprises contacting the nucleic acid sample with a mixture of oligonucleotides comprising (i) a set of

20     labeled primers each of whose nucleotide sequences comprises a target subsequence and (ii) an unlabeled primer whose sequence comprises one of the target subsequences identified in (i) followed by at least one additional nucleotide.

5.     The method of claim 3 wherein the extending of the sequence information comprises

25     contacting the nucleic acid sample with a mixture of oligonucleotides comprising (i) a set comprising a first unlabeled primer and a second unlabeled primer each of whose nucleotide sequence comprises a target subsequence and (ii) a set comprising a labeled third primer whose sequence comprises the subsequence of the first unlabeled primer and a labeled fourth primer whose sequence comprises the subsequence of the second unlabeled primer extended by at least

30     one nucleotide.

6.      The method of claim 1 wherein at least one of said generated signals corresponds to a sequence having a size and target subsequence of a sequence present in said sequence database.

7.      The method of claim 1 wherein said method additionally includes

recovering a fragment of a nucleic acid in the sample which generates said signal;

sequencing said fragment to determine at least a partial sequence for said fragment; and

verifying that said sample comprises a nucleic acid having a sequence comprising at least a portion of said determined sequence.

8.      The method of claim 1 wherein said plurality of nucleic acids are DNA.

9.      The method of claim 8, wherein said probing comprises:

digesting the sample with one or more restriction endonucleases, said restriction endonucleases having recognition sites that are said target subsequences and leaving single-stranded nucleotide overhangs on the digested ends;

hybridizing double-stranded adapter nucleic acids with the digested sample fragments, said adapter nucleic acids having an end complementary to one of said single-stranded overhangs; and

ligating the complementary ends end the of adapter nucleic acids to the complementary 5'-end of a strand of the digested sample fragments to form ligated nucleic acid fragments.

10.     The method of claim 7, wherein said plurality of nucleic acids are RNA.

11.     A method for extending the sequence in a length-subsequence combination of one or more nucleic acids in a sample comprising a plurality of nucleic acids having different nucleotide sequences, said method comprising:

33

(a) probing said sample with one or more recognition means wherein each recognition means recognizes a different target nucleotide subsequence or a different set of target nucleotide subsequences to provide one or more targeted nucleic acids;

5     (b) generating one or more first signals from said sample probed by said recognition means, each generated first signal arising from a targeted nucleic acid in said sample and comprising a representation of (*i*) the length between occurrences of target subsequences in said targeted nucleic acid, and (*ii*) the identities of said target subsequences in said targeted nucleic acid or identities of said target subsequences among which are included the target subsequences

10     in said targeted nucleic acid;

(c) selecting one or more targeted nucleic acids based on their corresponding first signals; and

(d) extending sequence information from one or more target subsequences in said targeted nucleic acid by one or more nucleotides providing one or more extended subsequences

15     under conditions that generate one or more second signals arising from selected targeted nucleic acid in said sample at least one of whose subsequences has been extended, wherein said second signal comprises a representation of (*i*) the length between occurrences of target subsequences, at least one of which has been extended, in said nucleic acid, and (*ii*) the identities of said target subsequences, at least one of which has been extended, in said selected targeted nucleic acid or

20     identities of said target subsequences, at least one of which has been extended, among which are included the target subsequences in said selected targeted nucleic acid;

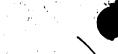whereby a matched nucleic acid in said sample has an extended sequence in said length-subsequence combination.

25     12.     The method of claim 11 wherein said second generated signal is a negative oligo-competition signal.

13.     The method of claim 11 wherein said second generated signal is a positive oligo-competition signal.

30

14.    The method of claim 12 wherein the extending of the sequence information comprises contacting the nucleic acid sample with a mixture of oligonucleotides comprising (i) a set of labeled primers each of whose nucleotide sequences comprises a target subsequence and (ii) an unlabelled primer whose sequence comprises one of the target subsequences identified in (i) followed by at least one additional nucleotide.

15.    The method of claim 13 wherein the extending of the sequence information comprises contacting the nucleic acid sample with a mixture of oligonucleotides comprising (i) a set comprising a first unlabeled primer and a second unlabeled primer each of whose nucleotide sequence comprises a target subsequence and (ii) a set comprising a labeled third primer whose sequence comprises the subsequence of the first unlabeled primer and a labeled fourth primer whose sequence comprises the subsequence of the second unlabeled primer extended by at least one nucleotide.